Unlocking New Capabilities for Azure Digital Twins with Real-Time Analytics

By Dr. William Bain, Founder & CEO | ScaleOut Software

HIGHLIGHTS

▶ The digital twin model can dramatically simplify applications that implement real-time analytics while providing much faster alerting and feedback.

▶ ScaleOut Software's scalable, in-memory computing platform seamlessly adds real-time analytics to the Azure Digital Twins cloud service.

▶ This combined solution incorporates machine learning with message processing to bring powerful new capabilities to a wide range of applications.

THE NEED FOR REAL-TIME ANALYTICS WITH DIGITAL TWINS

In countless applications that track live systems, real-time analytics plays a key role in identifying problems (or finding opportunities) and responding fast enough to make a difference. Consider a software telematics application that tracks a nationwide fleet of trucks to ensure timely deliveries. Dispatchers receive telemetry from trucks every few seconds detailing location, speed, lateral acceleration, engine parameters, and cargo viability. In a classic needle-and-haystack scenario, dispatchers must continuously sift through telemetry from thousands of trucks to spot issues, such as lost or fatigued drivers, engines requiring maintenance, or unreliable cargo refrigeration. They must intervene quickly to keep the supply chain running smoothly. Real-time analytics

ScaleOut Software

can help dispatchers tackle this seemingly impossible task by automatically sifting through telemetry as it arrives, analyzing it for anomalies needing attention, and alerting dispatchers when conditions warrant.

By using a process of divide and conquer, digital twins can dramatically simplify the construction of applications that implement real-time analytics for telematics or other applications. A digital twin for each truck can track that truck's parameters (for example, maintenance and driver history) and its dynamic state (location, speed, engine and cargo condition, etc.). The digital twin can analyze telemetry from the truck to update this state information and generate alerts when needed. It can encapsulate analytics code or use machine learning techniques to look for anomalies. Running simultaneously, thousands of digital twins can track all the trucks in a fleet to keep dispatchers informed while reducing their workload.

Applying the digital twin model to real-time analytics expands its range of uses from its traditional home in product lifecycle management and infrastructure tracking to managing time-critical, live systems with many data sources. Examples include preventive maintenance, health-device tracking, logistics, physical and cyber security, IoT for smart cities, ecommerce shopping, financial services, and many others. But how can we integrate real-time analytics with digital twins and ensure high performance combined with straightforward application development?

MESSAGE PROCESSING WITH AZURE DIGITAL TWINS

Microsoft's Azure Digital Twins provides a compelling platform for creating digital twin models with a rich set of features for describing their contents, including properties, components, inheritance, and more. The Azure Digital Twins Explorer GUI tool lets users view digital twin models and instances, as well as their relationships.

Azure digital twins can host dynamic properties that track the current state of physical data sources. Users can create serverless functions using Azure Functions to ingest messages generated by data sources and delivered to digital twins via Azure IoT Hub (or other message hubs). These functions update the properties of Azure digital twins using APIs provided for this purpose. Here's a redrawn tutorial example that shows how Azure functions can process messages from a thermostat and update both its digital twin and a parent digital twin that models the room in which the thermostat is located. Note that the first Azure function's update triggers the Azure Event Grid to run a second function that updates the room's property:
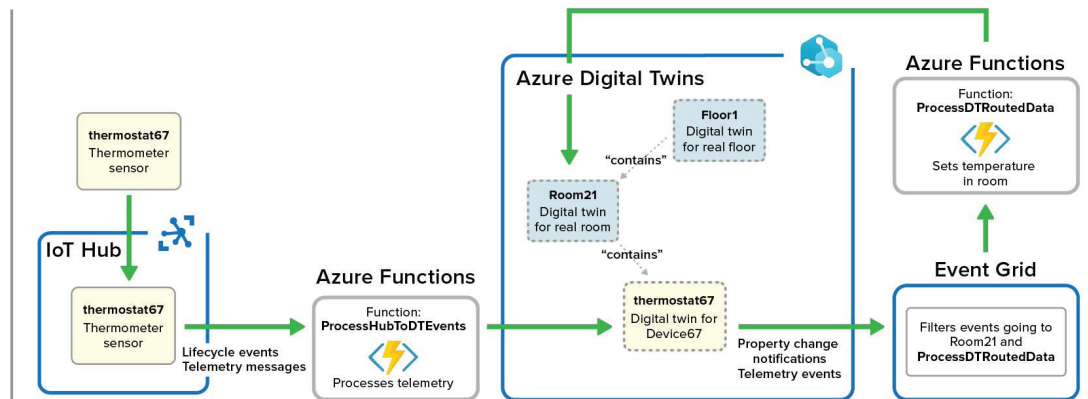
The challenge in using serverless functions to process messages and perform real-time analytics is that they add overhead and complexity. By their nature, serverless functions are stateless and must obtain their state from external services; this adds latency. In addition, they are subject to scheduling and authentication overheads on each invocation, and this adds delays that limit scalability. The use of multiple serverless functions and associated mechanisms, such as Event Grid topics and routes, also adds complexity in developing analytics code.

ADDING REAL-TIME ANALYTICS USING IN-MEMORY COMPUTING

Integrating an in-memory computing platform with the Azure Digital Twins infrastructure addresses both of the challenges. This technology runs on a cluster of virtual servers and hosts application-defined software objects in memory for fast access along with a software-based compute engine that can run application-defined methods with extremely low latency. By storing each Azure digital twin instance's properties in memory and routing incoming messages to an in-memory method for processing, both latency and complexity can be dramatically reduced, and real-time analytics can be scaled to handle thousands or even millions of data sources.

ScaleOut Software's newly announced Azure Digital Twins Integration does just this. It integrates the ScaleOut Digital Twin Streaming Service™, an in-memory computing platform running on Microsoft Azure (or on premises), with the Azure Digital Twins service to provide real-time streaming analytics. It accelerates message processing using in-memory computing to ensure fast, scalable performance while simultaneously streamlining the programming model.

The ScaleOut Azure Digital Twins Integration creates a component within an Azure Digital Twin model in which it hosts "real-time" properties for each digital twin instance of the model. These properties track dynamic changes to the

instance's physical data source and provide context for real-time analytics.

To implement real-time analytics code, application developers create a message-processing method for an Azure digital twin model. This method can be written in C# or Java, using an intuitive rules-based language, or by configuring machine learning (ML) algorithms implemented by Microsoft's ML.NET library. It makes use of each instance's real-time properties, which it stores in a memory-based object called a real-time digital twin, and the in-memory compute engine automatically persists these properties in the Azure digital twin instance.

Here's a diagram that illustrates how real-time digital twins integrate with Azure digital twins to provide real-time streaming analytics:
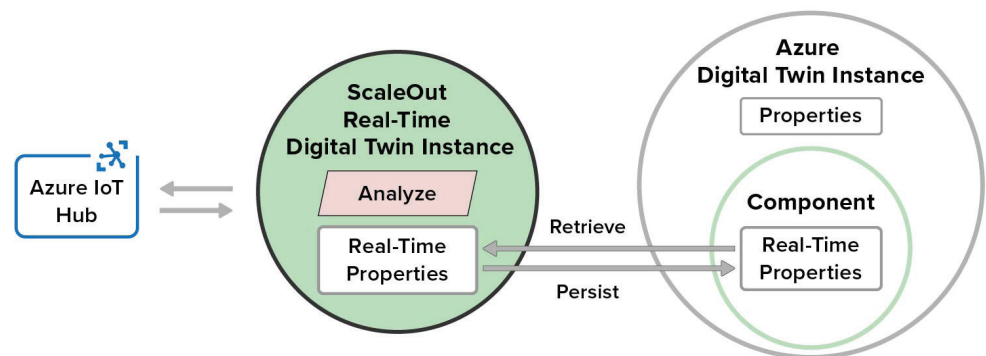
This diagram shows how each real-time digital twin instance maintains in-memory properties, which it retrieves when deployed, and automatically persists these properties in its corresponding Azure digital twin instance. The real-time digital twin connects to Azure IoT Hub or other message source to receive and then analyze incoming messages from its corresponding data source. Fast, in-memory processing provides sub-millisecond access to real-time properties and completes message processing with minimal latency. It also avoids repeated authentication delays every time a message is processed by authenticating once with the Azure Digital Twins service at startup.

All real-time analytics performed during message processing can run within a single in-memory method that has full access to the digital twin instance's properties. This code also can access and update properties in other Azure digital twin instances. These features simplify design by avoiding the need to split functionality across multiple serverless functions and by providing a straightforward, object-oriented design framework with advanced, built-in capabilities, such as ML.

To further accelerate development, ScaleOut provides tools that automatically

generate Azure digital twin model definitions for real-time properties. These model definitions can be used either to create new digital twin models or to add a real-time component to an existing model. Users just need to upload the model definitions to the Azure Digital Twins service.

Here's how the tutorial example for the thermostat would be implemented using ScaleOut's Azure Digital Twins Integration:
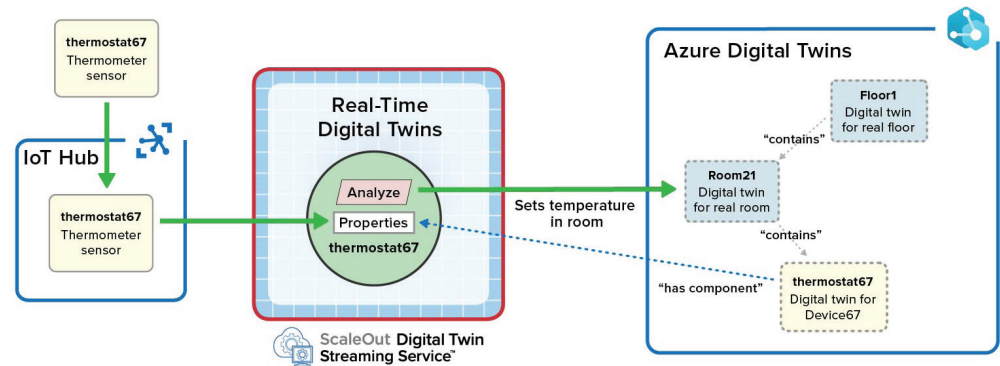
Note that the ScaleOut Digital Twins Streaming Service takes responsibility for ingesting messages from Azure IoT Hub and for invoking analytics code for the data source's incoming messages. Multiple, pipelined connections with Azure IoT Hub ensure high throughput. Also note that the two serverless functions and use of Event Grid have been eliminated since the in-memory method handles both message processing and updates to the parent object (Room 21).

Combining the ScaleOut Digital Twin Streaming Service with Azure Digital Twins gives users the power of in-memory computing for real-time analytics while leveraging the full spectrum of Azure services and tools, as illustrated below for the thermostat example:
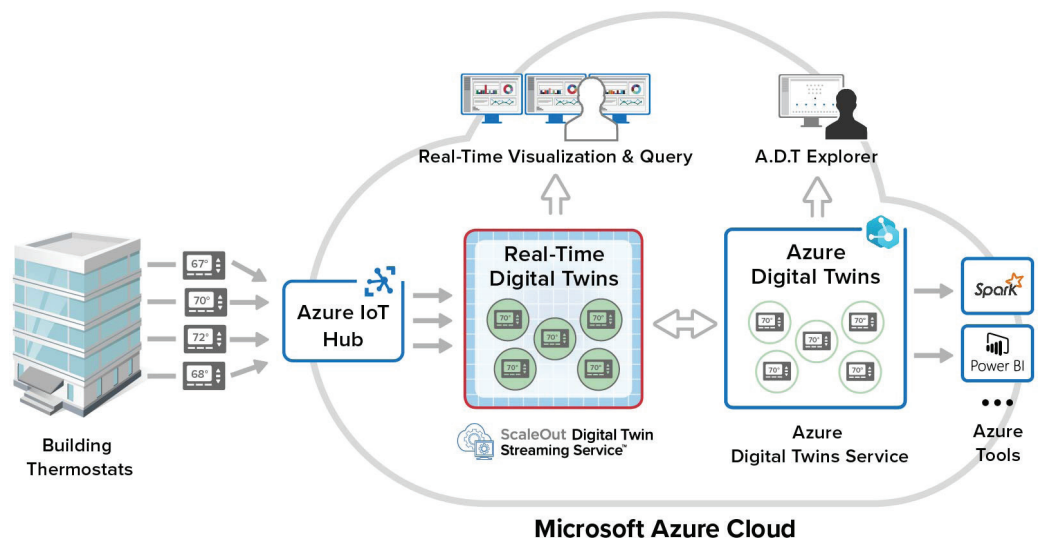
Users can view real-time properties with the Azure Digital Twins Explorer tool and track changes due to message processing. They also can take advantage of Azure's ecosystem of big data analytics tools like Spark to perform batch processing. ScaleOut's real-time data aggregation, continuous query, and visualization tools for real-time properties enable second-by-second tracking of live systems that boosts situational awareness for users.

EXAMPLE OF REAL-TIME ANALYTICS WITH AZURE DIGITAL TWINS

Incorporating real-time analytics using ScaleOut's Azure Digital Twins Integration unlocks a wide array of applications for Azure Digital Twins. For example, here's how the telematics software application discussed above could be implemented:
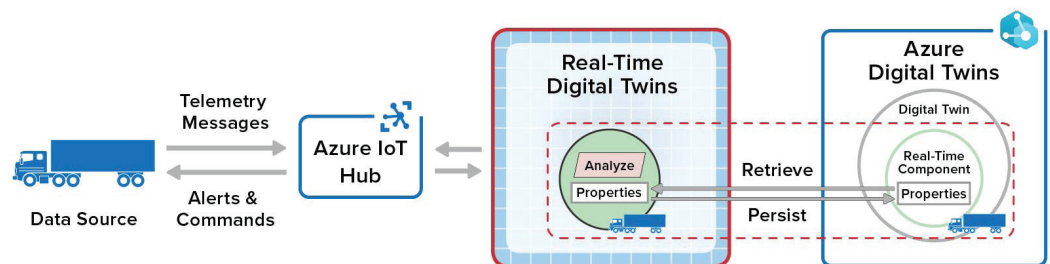
Each truck has a corresponding Azure digital twin which tracks its properties including a subset of real-time properties held in a component of each instance. When telemetry messages flow in to Azure IoT Hub, they are processed and analyzed by ScaleOut's in-memory computing platform using a real-time digital twin that holds a truck's real-time properties in memory for fast access and a message-processing method that analyzes telemetry changes, updates properties, and signals alerts when needed.

Real-time analytics can run ML algorithms that continuously examine telemetry, such as engine parameters, to detect anomalies and signal alerts. Digital twin analytics, combined with data aggregation and visualization powered by the in-memory platform, enable dispatchers to quickly spot emerging issues and take corrective action in a timely manner.

SUMMING UP

Digital twins offer a powerful means to model and visualize a population of physical devices. Adding real-time analytics to digital twins extends their reach into live, production systems that perform time-sensitive functions. By enabling managers to continuously examine telemetry from thousands or even millions of data sources and immediately identify emerging issues, they can avoid costly

problems and capture elusive opportunities.

Azure Digital Twins has emerged as a compelling platform for hosting digital twin models. With the integration of in-memory computing technology using the ScaleOut Digital Twin Streaming Service, Azure Digital Twins gains the ability to analyze incoming telemetry with low latency, high scalability, and a straightforward development model. The combination of these two technologies has the potential to unlock a wide range of important new use cases for digital twins.

ABOUT THE AUTHOR

Dr. William L. Bain is founder & CEO of ScaleOut Software, Inc. and holds a Ph.D. in electrical engineering from Rice University. Bill has contributed to advancements in parallel computing at Bell Labs research, Intel and Microsoft, and holds several patents in computer architecture and distributed computing. Bill founded and ran three start-up companies prior to founding ScaleOut Software in 2003. His previous company, Valence Research, developed and distributed Web load-balancing software, and was acquired by Microsoft Corporation. As a member of the screening committee for the Seattle-based Alliance of Angels, Bill is actively involved in entrepreneurship and the angel community.

ABOUT SCALEOUT SOFTWARE

Founded in 2003, ScaleOut Software develops leading-edge software that delivers scalable, highly available, in-memory computing and streaming analytics technologies to a wide range of industries. ScaleOut Software's in-memory computing platform enables operational intelligence by storing, updating, and analyzing fast-changing, live data so that businesses can capture perishable opportunities before the moment is lost.

ScaleOut Software offers a full portfolio of in-memory computing products and technologies to deliver operational intelligence for your organization. Please visit www.scaleoutsoftware.com for more details on our technology and product portfolio for mission critical enterprise applications, including the ScaleOut Digital Twin Streaming Service™, an award-winning Azure-based cloud service and on-premises platform for hosting streaming analytics using the digital twin model. In addition, ScaleOut Software offers ScaleOut StateServer®, a linearly scalable, highly available in-memory data grid and ScaleOut StateServer® Pro, a data-parallel compute engine for blazingly fast, data-parallel computation on live, fast-changing data.

WINNER OF THE
2022 IoT ANALYTICS
INNOVATION AWARD